DOI: 10.24850/j-tyca-14-03-01

Artículos

Análisis de frecuencias conjunto de gastos pico y volúmenes de crecientes con marginales Gumbel

Joint frequency analysis of peak flow and volumes of floods with Gumbel marginals

Daniel Francisco Campos-Aranda[1], ORCID: https://orcid.org/0000-0002-9876-3967

[1]Universidad Autónoma de San Luis Potosí, San Luis Potosí, México, campos_aranda@hotmail.com

Autor para correspondencia: Daniel Francisco Campos-Aranda, campos_aranda@hotmail.com

Resumen

Desde hace dos décadas, la estimación de las *crecientes de diseño* de los embalses se aborda con el enfoque multivariado más simple: el *bivariado*. Lo anterior se aceptó pues se demostró que los embalses no son sensibles al tiempo al gasto máximo, y que tal gasto y volumen están correlacionados entre ellos y este último con la duración total del

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

hidrograma de la creciente. En este estudio se ajustó la distribución Gumbel bivariada o *modelo logístico* a los 61 datos anuales de gasto pico y volumen de las crecientes de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*) en el río Mayo del estado de Sonora, México. Este proceso abarca las *ocho* etapas siguientes: (1) selección y prueba de los registros por procesar; (2) verificación de la aleatoriedad de los registros anuales; (3) aceptación de las funciones marginales Gumbel; (4) estimación de las probabilidades empíricas conjuntas; (5) validación del modelo Logístico; (6) verificación de las restricciones de probabilidad; (7) estimación de eventos de diseño, gasto pico y volumen, univariados híbridos, y (8) estimación de eventos de diseño conjuntos. En la etapa 1 primero se hace una selección subjetiva y después se verifica con el Test PPCC. La etapa 2 se realiza con base en el Test de Wald-Wolfowitz. Las etapas 3 y 5 utilizan el Test de Kolmogórov-Smirnov. En la etapa 7 se definen gastos de diseño, y se obtienen volúmenes por regresión y probabilidad condicional. En contraste, en la etapa 8 se obtienen diversos eventos de gasto pico y volumen que pertenecen al *subgrupo de parejas críticas* en las gráficas del periodo de retorno conjunto $T'(Q,V)$. Por último, se formulan las Conclusiones, las cuales destacan las ventajas del análisis de frecuencias conjunto bivariado y la sencillez de aplicación y prueba del modelo Logístico.

Palabras clave: crecientes de diseño, distribuciones Gumbel bivariadas, distribuciones condicionales del modelo Logístico, probabilidades empíricas conjuntas, validación del modelo Logístico, periodos de retorno univariados híbridos, periodos de retorno conjuntos.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Abstract

For two decades, the estimation of *Design Floods* of reservoirs has been addressed with the simplest multivariate approach, the *bivariate*. This has been accepted because it was proven that the reservoirs are not time sensitive to maximum flow, moreover, that such flow and volume are correlated with each other and the latter, with the total duration of the flood hydrograph. In this study, the bivariate Gumbel distribution or *Logistics model* was adjusted to the 61 annual data of peak flow and volume of floods entering the Adolfo Ruiz Cortines (*Mocúzari*) dam in the Río Mayo of the state of Sonora, Mexico. This process comprehends the following *eight* stages: (1) selection and testing of records to be processed; (2) verification of the randomness of the annual records; (3) acceptance of Gumbel marginal functions; (4) estimation of the joint empirical probabilities; (5) validation of the Logistic model; (6) verification of probability constraints; (7) estimation of design events, peak flow and volume, hybrid univariates, and (8) estimation of joint design events. In stage 1, first a subjective selection is made and then it is verified with the PPCC Test. Stage 2 is carried out based on the Wald-Wolfowitz Test. Stages 3 and 5 use the Kolmogórov-Smirnov Test. In stage 7, design flows are defined, and volumes are obtained by regression and conditional probability. In contrast, in stage 8, several peak flow and volume events are obtained, belonging to the *subgroup of critical pairs*, in the graphs of the joint return period $T'(Q,V)$. Towards the last part of this work, conclusions are formulated, which highlight the advantages of the bivariate joint frequency analysis and the simplicity of application and testing of the Logistic model.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Keywords: Design floods, bivariate Gumbel distributions, conditional distributions of the Logistic model, joint empirical probabilities, validation of the Logistic model, hybrid univariate return periods, joint return periods.

Recibido: 07/01/2021

Aceptado: 21/10/2021

Introducción

Generalidades

De manera preponderante, la porción centro-sur de la república mexicana se localiza en la zona de influencia de los huracanes que se originan en los océanos Atlántico y Pacífico. Por otra parte, su porción norte la afectan los frentes fríos. Ambos fenómenos meteorológicos generan lluvias de gran magnitud, que producen *crecientes* o *avenidas máximas*, las cuales inundan extensas regiones y ponen en peligro a las obras hidráulicas. Ante tal panorama que se repite cada año en diversas zonas del país

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

resulta evidente la necesidad de estudiar las crecientes para formular medidas de protección y mitigación de los daños sociales, ambientales y económicos que producen tales inundaciones (Aldama, 2000; Aldama, Ramírez, Aparicio, Mejía-Zermeño, & Ortega-Gil, 2006; Domínguez & Arganis, 2012).

En relación con el dimensionamiento hidrológico de los diques o muros de protección, puentes, rectificaciones y canalizaciones, es necesario definir las *crecientes de diseño* (CD), o gastos máximos asociados con bajas probabilidades de ser excedidos. La estimación de las CD se realiza mediante el *análisis de frecuencia de crecientes* (AFC), proceso que involucra representar los datos anuales de las crecientes por un modelo probabilístico o *función de distribución de probabilidades* (FDP), con lo cual se realizan las *predicciones* buscadas (Kite, 1977; Bobée & Ashkar, 1991; Hosking & Wallis, 1997; Rao & Hamed, 2000; Meylan, Favre, & Musy, 2012; Stedinger, 2017).

Desde hace dos décadas, las crecientes de diseño de los embalses se analizan como *eventos multivariados*, pues las características que definen su *hidrograma* están correlacionadas. El enfoque más simple, el *bivariado*, se justifica por la escasa sensibilidad que tienen los almacenamientos al lapso en que se llega al gasto pico (Aldama, 2000), y por la correlación que guarda tal gasto máximo con el volumen y el volumen con la duración total del hidrograma (Goel, Seth, & Chandra, 1998; Yue, 1999; Yue, Ouarda, Bobée, Legendre, & Bruneau, 1999; Yue, 2000c; Yue & Rasmussen, 2002).

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Ramírez-Orozco y Aldama (2000), Escalante-Sandoval y Reyes-Chávez (2002), Volpi y Fiori (2012), Requena, Mediero y Garrote (2013), y Vogel y Castellarin (2017) destacan que el AFC bivariado conduce a una infinidad de combinaciones de gasto pico y volumen para una probabilidad de excedencia *conjunta* adoptada. Lo anterior implica que para un mismo *periodo de retorno conjunto* existen muchas crecientes o *hidrogramas* que producirán distintos efectos en el embalse que se diseña o revisa, adoptando, lógicamente, el que genera las condiciones más críticas o severas en su almacenamiento y vertedor.

Las tormentas severas son otros eventos hidrológicos que se han analizado con un enfoque bivariado, cuya intensidad máxima diaria está correlacionada con la cantidad de lluvia ocurrida durante los días de tal evento (Yue, 2000a; Yue, 2000b).

Objetivos

El *objetivo* fundamental del estudio consistió en exponer el ajuste del modelo probabilístico bivariado más simple, por su naturaleza explícita, con marginales de tipo Gumbel, denominado *modelo Logístico*. Lo anterior se desarrolló en los seis *objetivos* siguientes: (1) para tal distribución bivariada se citan las ecuaciones de su FDP conjunta y de probabilidad condicional; (2) se expone la teoría que respalda el tema de los periodos

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)
Tecnología y ciencias del agua, ISSN 2007-2422, 14(3), 01-55. DOI: 10.24850/j-tyca-14-03-01

de retorno univariados con regresión y condicionales, y conjuntos; (3) se describe cómo se estiman las probabilidades empíricas univariadas y bivariadas; (4) se muestra cómo se verifican las funciones marginales y valida el modelo Logístico; (5) se concluye con la selección y prueba de los registros por procesar, y (6) la teoría expuesta se aplica a los 61 gastos pico y volúmenes anuales de las crecientes de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*) sobre el río Mayo en el estado de Sonora, México.

Teoría operativa

Descripción del análisis de frecuencias conjunto

Se inicia seleccionando registros aleatorios con estacionariedad de gasto pico (*Qp*) y volumen (*V*) anuales de crecientes que no muestren valores dispersos (*outliers*) ni procedan de tipos diferentes de crecientes. Enseguida se verifica que ambos registros pueden ser representados probabilísticamente por distribuciones Gumbel mediante el test PPCC.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Después se ajusta la distribución Gumbel a cada registro de *Qp* y *V*, y se ratifica su aceptación de manera gráfica y por medio del test de Kolmogórov-Smirnov (KS). Se prosigue con la estimación de la correlación entre las variables *Qp* y *V* para poder ajustar el modelo logístico y estimar las probabilidades conjuntas teóricas, las cuales se contrastan con las empíricas nuevamente de manera gráfica y con el test de KS.

Por último, se aborda la estimación de los periodos de retorno univariados híbridos y después la de los conjuntos de tipo AND. Haciendo variados cálculos de este último, se forma la gráfica que incluye una curva para cada periodo de retorno de diseño. En la parte curva de cada gráfica se obtienen los eventos críticos de diseño.

Ajuste de las distribuciones Gumbel bivariadas

Existen dos distribuciones bivariadas de valores extremos *explícitas*. La primera se denomina Gumbel Mixta (GM) y la segunda Gumbel Logística (GL), ambas fueron propuestas por Emil Julius Gumbel a inicios de la década de 1960. Las FDP de ambas distribuciones bivariadas son (Yue *et al*., 1999; Ramírez-Orozco & Aldama, 2000; Yue & Rasmussen, 2002; Yue & Wang, 2004):

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$F_{GM}(x,y) = F_X(x) \cdot F_Y(y) \cdot \exp\left\{-\theta\left[\frac{1}{\ln F_X(x)} + \frac{1}{\ln F_Y(y)}\right]^{-1}\right\} \; (0 \leq \theta \leq 1) \qquad (1)$$

$$F_{GL}(x,y) = \exp\{-[(-\ln F_X(x))^m + (-\ln F_Y(y))^m]^{1/m}\} \; (m \geq 1) \qquad (2)$$

donde $F_X(x)$ y $F_Y(y)$ son las FDP *marginales* de las variables aleatorias X y Y, cuyas expresiones son:

$$F_X(x) = \exp\left[-\exp\left(-\frac{x-u_x}{\alpha_x}\right)\right] \; (x \geq 0) \qquad (3)$$

$$F_Y(y) = \exp\left[-\exp\left(-\frac{y-u_y}{\alpha_y}\right)\right] \; (y \geq 0) \qquad (4)$$

siendo u y a los parámetros de ubicación y escala de cada distribución Gumbel, también designada *doble exponencial*; sus expresiones de acuerdo con el método de momentos en función de la media (M) y desviación estándar (S) insesgada de la muestra son (Kite, 1977; Rao & Hamed, 2000):

$$\alpha = 0.7797 \cdot S \qquad (5)$$

$$u = M - 0.5772 \cdot \alpha \qquad (6)$$

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Las ecuaciones (5) y (6) son aplicables en registros *X* y *Y* que tienen igual amplitud (*n*). Para procesar registros de tamaños diferentes se debe aplicar el método de ajuste por máxima verosimilitud expuesto por Escalante-Sandoval (2005), para obtener los cinco parámetros del modelo logístico (Ecuación (2)).

La solución inversa de las ecuaciones (3) y (4) permite la estimación de *predicciones* (x_p,y_p) asociadas con una probabilidad de no excedencia $p = F_X(x)$ o $p = F_Y(y)$ y son las siguientes:

$$x_p = u_x + \alpha_x \cdot \{-\ln[-\ln(p)]\} \tag{7}$$

$$y_p = u_y + \alpha_y \cdot \{-\ln[-\ln(p)]\} \tag{8}$$

En ambos modelos bivariados, θ y m son los *parámetros de asociación*, que describen la dependencia entre las dos variables aleatorias. Sus expresiones son:

$$\theta = 2 \cdot \left[1 - \cos\left(\pi \cdot \sqrt{\frac{\rho}{6}}\right)\right] \ (0 \leq \rho \leq 2/3) \tag{9}$$

$$m = \frac{1}{\sqrt{1-\rho}} \ (0 \leq \rho \leq 1) \tag{10}$$

siendo ρ el coeficiente de correlación de Pearson, cuya ecuación es:

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$\rho = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \cdot \sigma_Y} \quad (11)$$

en la cual (μ_X, σ_X) y (μ_Y, σ_Y) son la media y la desviación estándar *poblacionales* de las variables aleatorias *X* y *Y*.

Yue y Wang (2004) indican que cuando el coeficiente ρ varía entre cero y 2/3, ambos modelos bivariados (ecuaciones (1) y (2)) conducen a probabilidades conjuntas idénticas, y cuando ρ fluctúa de 2/3 a 1, solo el modelo logístico (GL) debe ser utilizado. Por lo anterior, en adelante únicamente se trabaja con tal modelo probabilístico.

Al respecto, Escalante-Sandoval (2005) emplea el modelo logístico a nivel regional para procesar crecientes y concluye que sus predicciones son menos sesgadas que las estimaciones univariadas. Además, indica que las probabilidades univariadas y conjunta del modelo logístico deben cumplir con la restricción siguiente:

$$F_X(x) \cdot F_Y(y) < F_{GL}(x, y) < min[F_X(x), F_Y(y)] \quad (12)$$

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Se definen de manera estadística ortodoxa para la variable *X* dado que *Y* = *y*, y similarmente para *Y*, dado que *X* = *x*. Sus expresiones son (Yue & Rasmussen, 2002):

$$F(x|Y=y) = F(x,y) \cdot [EMX + EMY]^{(1-m)/m} \cdot \exp(MY + EY) \quad (13)$$

$$F(y|X=x) = F(x,y) \cdot [EMX + EMY]^{(1-m)/m} \cdot \exp(MX + EX) \quad (14)$$

siendo:

$$EMX = \exp[-m(x-u_x)/\alpha_x] \quad (15)$$

$$EMY = \exp\left[-m\left(y-u_y\right)/\alpha_y\right] \quad (16)$$

$$MY = (1-m)(y-u_y)/\alpha_y \quad (17)$$

$$MX = (1-m)(x-u_x)/\alpha_x \quad (18)$$

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$EY = \exp\left[-(y - u_y)/\alpha_y\right] \tag{19}$$

$$EX = \exp[-(x - u_x)/\alpha_x] \tag{20}$$

Yue (2000b), y Yue y Rasmussen (2002) han señalado, para la práctica hidrológica, que puede ser de mayor interés la condición de *X* dado *Y* ≤ *y* y de *Y* dado *X* ≤ *x*. Tales distribuciones condicionales del modelo mixto han sido expuestas por Yue *et al*. (1999) y Yue (2000b), y para el modelo logístico por Yue y Rasmussen (2002). Por otra parte, Shiau (2003) ha desarrollado las *condicionales* del modelo logístico de *X* dado *Y* ≥ *y* y de *Y* dado *X* ≥ *x*.

Periodos de retorno univariados híbridos

Periodos de retorno univariados

El concepto clásico de probabilidad de un evento se define como el cociente del número de casos favorables (*ncf*) a tal evento entre el número de casos posibles (*ncp*) a dicho evento, por ello varía de 0 a 1. Debido al manejo anual de la variable *X*, la probabilidad de excedencia

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$F'_X(x)$ se corresponde con el recíproco del *periodo de retorno* (T_X) en años, ya que en cada año se tiene *ncf* =1 y *ncp* =T_X, esto es (Yue & Rasmussen, 2002; Shiau, 2003):

$$T_X = \frac{1}{F'_X(x)} = \frac{1}{1-F_X(x)} \tag{21}$$

En la expresión anterior, $F_X(x)$ es la probabilidad de no excedencia que se estima con la Ecuación (3), o la Ecuación (4) para T_Y.

Cuando se integran registros de crecientes con la técnica POT (*peaks-over-threshold*), se forman las llamadas series de duración parcial, que incluyen más datos que el número de años de registro y entonces de acuerdo con Shiau (2003), el numerador de la ecuación anterior se debe sustituir por $E(L) < 1$ o valor promedio del tiempo entre crecientes.

En el contexto bivariado, como el de las crecientes y con un enfoque estricto, los periodos de retorno univariados sólo son aplicables cuando existe una única variable aleatoria que es dominante o significativa en el criterio de diseño, o bien la dependencia entre las dos variables es baja o escasa (Shiau, 2003).

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Periodos de retorno condicionales

Se obtienen aplicado la Ecuación (21), previa sustitución de $F_X(x)$ por $F(x|Y=y)$ o por $F(y|X=x)$, es decir:

$$T(x|y) = \frac{1}{1-F(x|Y=y)} \tag{22}$$

$$T(y|x) = \frac{1}{1-F(y|X=x)} \tag{23}$$

Eventos de diseño obtenidos con regresión

Cuando es posible determinar la variable X dominante o de mayor influencia (*driving variable*) en el diseño hidrológico, entonces es posible seleccionar un periodo de retorno T_X y con base en la Ecuación (21) obtener el valor de diseño x_d, el cual se obtiene con la FDP marginal de X (Ecuación (7)). Después, a partir de una ecuación de regresión entre las variables X y Y, se aplica x_d como variable independiente y se obtiene la y_d buscada. Este enfoque no estima un periodo de retorno conjunto, pero aporta una estimación (x_d,y_d) que tiene un soporte estadístico (Serinaldi & Grimaldi, 2011; Gräler *et al*., 2013).

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Eventos de diseño de la distribución bivariada condicional

En este otro enfoque, el valor de x_d o predicción del periodo de retorno univariado de la variable dominante o prominente X se lleva a la distribución bivariada *condicional* definida como $F(y|X = x_d)$ para obtener el valor de y_d buscado, el cual genera el mismo periodo de retorno de la variable X. De nuevo, esta aproximación no conduce a un evento de diseño bivariado que tenga un periodo de retorno conjunto, pero tal enfoque ha sido sugerido y aplicado por varios autores (Gräler *et al*., 2013).

Periodos de retorno conjuntos

El primer *periodo de retorno conjunto* del evento (X, Y) se define bajo la condición OR y la Ecuación (21) de la manera siguiente (Goel *et al*., 1998; Yue, 2000b; Shiau, 2003):

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$T(x,y) = \frac{1}{1-F(x,y)} \text{ siendo } F(x,y) = P(X \leq x \text{ ó } Y \leq y) \quad (24)$$

En la expresión anterior, $F(x,y)$ es la probabilidad de no excedencia conjunta que se estima con la Ecuación (2). Este evento representa el caso en que los límites x o y, o ambos *pueden* ser excedidos ($X > x$, o $Y > y$; o bien, $X > x$ y $Y > y$).

El segundo *periodo de retorno conjunto* del evento (X, Y) está asociado al caso en que ambos límites *son* excedidos ($X > x$, $Y > y$) o condición AND; su ecuación es (Goel *et al*., 1998; Yue, 2000b; Shiau, 2003):

$$T'(x,y) = \frac{1}{F'(x,y)} = \frac{1}{1+F(x,y)-F_X(x)-F_Y(y)} \quad (25)$$

Aldama (2000) obtiene la expresión $F'(x,y)$ de la probabilidad conjunta de excedencia mediante un razonamiento de probabilidades lógico y simple aplicado en el plano cartesiano. En cambio, Yue y Rasmussen (2002) recurren al plano cartesiano para definir numéricamente un evento bivariado (X, Y), que puede ocurrir en alguno de los cuatro cuadrantes y entonces obtienen la ecuación de la probabilidad conjunta de excedencia (denominador de la Ecuación (25)).

Yue y Rasmussen (2002) también establecen las siguientes relaciones entre los periodos de retorno univariados y los periodos de retorno conjuntos:

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$T(x,y) \leq \text{mínimo}\ (T_X, T_Y) \tag{26}$$

$$T'(x,y) \geq \text{máximo}\ (T_x, T_Y) \tag{27}$$

Diversos autores (Yue, 2000b; Yue & Rasmussen, 2002; Shiau, 2003; Yue & Wang, 2004; Requena *et al*., 2013; Vogel & Castellarin, 2017) han mostrado las gráficas de los dos periodos de retorno conjuntos y han discutido sus diferencias. En la Figura 1 se muestra la gráfica del periodo de retorno conjunto *T′*(*Q*,*V*), construida con los datos de la aplicación numérica, la cual se describe posteriormente.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

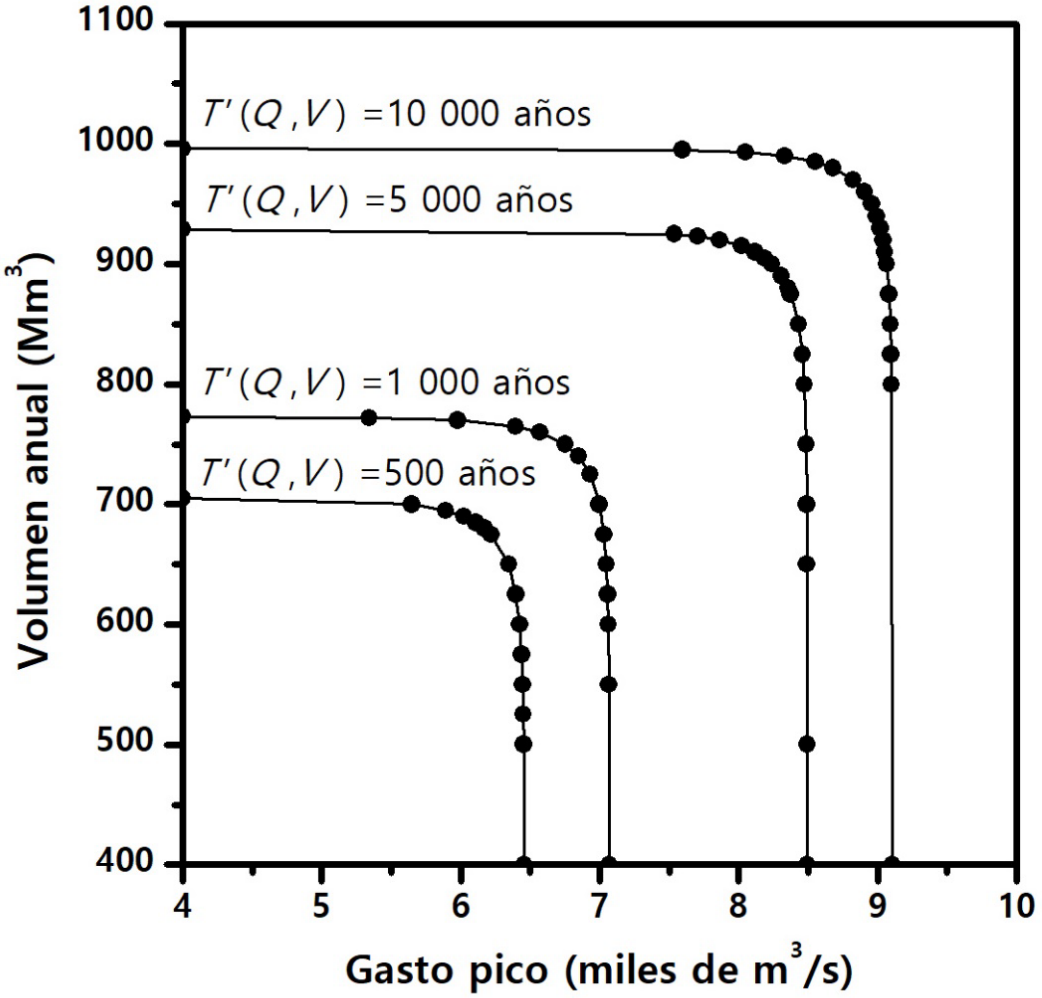

Figura 1. Gráficas de los cuatro periodos de retorno conjunto *T′*(*Q*,*V*) de diseño, del gasto pico y volumen anual de las crecientes de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*), México.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Recientemente, Salas, Obeysekera y Vogel (2018) han designado al periodo de retorno como el *tiempo de espera* entre la ocurrencia de una creciente que excede a la avenida de diseño. Este concepto aplicado al caso bivariado implica la estimación del *periodo de retorno conjunto* con base en $F'(x, y)$; es decir, con la ecuación (25).

Estimación de probabilidades empíricas

Las probabilidades empíricas univariadas y bivariadas se estimaron con base en la fórmula de Gringorten (Ecuación (28)), que de acuerdo con Yue *et al.* (1999), Yue (2000b), Yue y Rasmussen (2002), y Yue y Wang (2004) conduce a probabilidades de no excedencia (*p*) insesgadas para la distribución Gumbel o de valores extremos I. Tal fórmula es:

$$p = \frac{i-0.44}{n+0.12} \tag{28}$$

en la cual *i* es el número de cada dato cuando están ordenados de manera progresiva, y *n* es el número total de ellos, o amplitud del registro procesado.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Para la estimación de las probabilidades empíricas bivariadas se siguió el mismo principio que aplica para la Ecuación (28), pero se trabajó en el plano bidimensional, con los datos ordenados en forma progresiva, los gasto pico (*Qp*) en los renglones y los volúmenes (*V*) en las columnas. El plano formado es un cuadrado de *n* por *n* casillas, con *n* casillas en su diagonal principal, cuando el número de orden del renglón es igual al de la columna. Después, cada pareja de datos anual (*Qp* y *V*) se localiza en el plano bidimensional, y la casilla definida por la intersección del renglón y columna se identifica con el número *i* que corresponde al año histórico dibujado.

Cuando las *n* parejas de datos están dibujadas, se busca el año 1 y se define un área rectangular o cuadrada de valores menores de *Qp* y de *V*, cuyo *conteo* de casillas numeradas dentro es NM_1, o combinaciones de *Qp* y *V* menores. Calculados los *n* valores de NM_i, se aplica la fórmula de posición gráfica de Gringorten (Stedinger, 2017) para calcular la probabilidad empírica conjunta o bivariada:

$$F(x,y) = P(Q \leq q, V \leq v) = \frac{NM_i - 0.44}{n + 0.12} \tag{29}$$

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Validación del modelo logístico

Yue (2000a) indica que la relación entre las probabilidades conjuntas teóricas y las empíricas del gasto pico y el volumen permiten definir la validez de la distribución conjunta propuesta. La forma más simple de representarlas consiste en llevar al eje de las abscisas la probabilidad de no excedencia empírica y en el eje de las ordenadas la teórica; lógicamente, cada pareja de datos define un punto que coincide o se aleja de la recta a 45°. La inspección de la gráfica descrita y el valor del coeficiente de correlación —en estos casos, superior a 0.98— ratifican la validez del modelo probabilístico conjunto utilizado.

Yue (2000b), y Yue y Rasmussen (2002) aplican el test de Kolmogórov-Smirnov con un nivel de significancia (*a*) del 5 % para aceptar o rechazar la máxima diferencia (*dif*) entre las probabilidades conjuntas. Para evaluar la estadística (D_n) del test, la cual es función del número de datos (*n*), se utilizó la expresión que exponen Meylan *et al*. (2012), para *a* = 5 %:

$$D_n = \frac{1.358}{\sqrt{n}} \tag{30}$$

Si *dif* es menor que D_n se acepta el modelo probabilístico conjunto o distribución Gumbel bivariada (modelo logístico).

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Selección del registro por procesar

Aldama *et al*. (2006) exponen 16 registros de gasto pico y volumen anual de las crecientes de entrada a 15 embalses y uno en proyecto. A tales registros se les ajustó la distribución General de Valores Extremos mediante el método de los momentos L (Hosking & Wallis, 1997). Se buscó la pareja de registros de *Qp* y *V* con los menores parámetros de forma (*k*), lo cual implica que pueden ajustarse a una distribución Gumbel (k = 0). Se encontró, para las entradas a la presa La Boquilla sobre el río Conchos del estado de Chihuahua, -0.200 y -0.125 como valores más bajos de *k* para el *Qp* y *V*, respectivamente.

Como tales registros ya fueron procesados por el autor con la distribución Normal bivariada, se adoptaron los siguientes con valores de *k* de -0.259 y -0.138 para las 61 crecientes de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*) sobre el río Mayo en el estado de Sonora, México. Su área de cuenca es de 10 719 km^2 y sus datos de *Qp* y *V* se presentan en la Tabla 1, en cuyo final se exponen sus parámetros estadísticos: media (*M*), desviación estándar (*S*) y coeficiente de correlación lineal (ρ).

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Tabla 1. Gastos pico y volúmenes anuales de las crecientes de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*), México (Aldama *et al*., 2006).

Núm.	Año	*Qp* (m^3/s)	*V* (Mm^3)	No.	Año	*Qp* (m^3/s)	*V* (Mm^3)
1	1941	602.00	64.55	33	1973	1 421.40	152.92
2	1942	774.90	47.69	34	1974	1 262.10	119.09
3	1943	880.00	24.43	35	1975	666.60	74.89
4	1944	673.60	23.39	36	1976	950.50	96.05
5	1945	392.25	16.30	37	1977	1 028.40	113.97
6	1946	1 159.25	87.77	38	1978	2 302.50	262.73
7	1947	319.35	10.92	39	1979	1 751.70	191.82
8	1948	1 212.00	72.99	40	1980	1 217.30	159.71
9	1949	6 390.00	636.47	41	1981	1 546.70	188.50
10	1950	437.0	37.99	42	1982	1 418.80	165.50
11	1951	400.00	4.64	43	1983	2 794.70	383.90
12	1952	573.78	45.21	44	1984	2 344.30	278.50
13	1953	368.00	10.40	45	1985	1 473.30	231.43
14	1954	273.00	65.93	46	1986	1 046.10	309.86
15	1955	1 350.80	117.71	47	1987	578.50	52.18
16	1956	469.00	58.13	48	1988	1 443.90	175.22

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Núm.	Año	*Qp* (m^3/s)	*V* (Mm^3)	No.	Año	*Qp* (m^3/s)	*V* (Mm^3)
17	1957	1 772.30	146.97	49	1989	635.10	54.10
18	1958	1 009.60	111.71	50	1990	4 158.90	419.64
19	1959	1 661.00	182.99	51	1991	1 501.30	220.22
20	1960	3 955.00	381.46	52	1992	1 914.00	121.30
21	1961	1 517.40	238.58	53	1993	923.30	93.18
22	1962	893.20	96.62	54	1994	1 896.50	160.84
23	1963	1 534.40	216.07	55	1995	3 275.60	317.98
24	1964	1 154.20	335.24	56	1997	1 535.30	174.98
25	1965	1 755.80	157.57	57	1998	2 797.00	144.98
26	1966	1 755.80	147.12	58	1999	813.10	72.00
27	1967	1 782.20	256.19	59	2000	2 006.80	277.60
28	1968	808.30	116.01	60	2001	801.00	70.29
29	1969	813.10	72.00	61	2002	692.50	68.91
30	1970	967.00	94.66	*M*		1 477.266	158.709
31	1971	4 935.40	456.21	*S*		1 132.950	124.374
32	1972	1 326.40	195.06	*ρ*		0.8942	

Respecto a la procedencia de tales datos, Aldama *et al*. (2006) indican que los registros hidrométricos continuos, de naturaleza gráfica o tabular, permitieron identificar los datos de gasto y volumen de las

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

avenidas; por lo anterior, se deduce que ambos valores proceden de una creciente. Además, y por lo general, tanto el gasto pico como su volumen de escurrimiento respectivo incluyen el gasto base.

Pruebas de aceptación de las marginales Gumbel

Una primera prueba al registro seleccionado con base en el menor valor del parámetro de forma (k) de la distribución GVE consiste en dibujar tales datos en el papel de probabilidad extremo o de Gumbel-Powell (Chow, 1964), para verificar que definen una línea recta de manera aproximada. Ésta es una prueba gráfica poco objetiva.

Para evitar la subjetividad en la búsqueda de los registros de Qp y V que acepten la distribución Gumbel, se propone aplicar una prueba simple desarrollada por Stedinger, Vogel y Foufoula-Georgiou (1993), la cual se estableció al generar valores aleatorios con $k = 0$ con una distribución GVE ajustada con el método de momentos L; se observó que el estimador de k de tal muestra tiene media cero y varianza igual a $0.5633/n$. Con ello, los autores definen que cuando el valor absoluto de la estadística Z excede a la desviación Normal estándar de 1.645, se deduce que k es estadísticamente diferente de cero y por lo tanto se debe ajustar la distribución GVE y no la Gumbel. La estadística es:

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$Z = k\sqrt{n/0.5633} \tag{31}$$

Entonces, para la búsqueda de registros que acepten la distribución Gumbel, se les ajusta la distribución GVE con el método de momentos L (Hosking & Wallis, 1997) y se aplica la Ecuación (31) si el valor absoluto de Z no excede 1.645, tales datos se pueden representar por la distribución Gumbel.

La segunda prueba que se aplicó para la aceptación de las marginales Gumbel es la del coeficiente de correlación del gráfico de probabilidad, designada Test PPCC por sus siglas en inglés (Vogel, 1986). Inicia calculando los coeficientes de correlación lineal (Ecuación (11)) entre los datos *ordenados* de menor a mayor (x_i,y_i) y sus estimaciones con las ecuaciones (7) y (8) cuando se utiliza como probabilidad (p) la estimada con la fórmula de Gringorten (Ecuación (28)). Tales estimaciones, Vogel (1986) las llama predicciones *medianas* de orden (M_i), y su estadístico para el gasto pico (x_i) y el volumen (y_i) serán:

$$r_x = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(M_i-\bar{M})}{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(M_i-\bar{M})^2\right]^{1/2}} \tag{32}$$

$$r_y = \frac{\sum_{i=1}^{n}(y_i-\bar{x})(M_i-\bar{M})}{\left[\sum_{i=1}^{n}(y_i-\bar{y})^2 \sum_{i=1}^{n}(M_i-\bar{M})^2\right]^{1/2}} \tag{33}$$

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Vogel (1986) expone los puntos críticos igual a 1 000(1-r_c) en la Tabla 2, en función de la amplitud en años del registro probado (n) y del nivel de significancia de la prueba. A partir de tales valores se obtiene el coeficiente de correlación crítico (r_c). Si los valores r_x y de r_y de los registros que se prueban exceden a r_c, se acepta la distribución Gumbel.

Tabla 2. Puntos críticos (PC) igual a 1 000(1-r_c) del Test PPCC de Vogel (1986).

n	**Nivel de significancia**			***n***	**Nivel de significancia**		
	1 %	**5 %**	**10 %**		**1 %**	**5 %**	**10 %**
30	80.9	47.4	37.8	70	49.4	28.0	21.3
40	71.4	40.6	31.1	80	47.5	25.3	19.6
50	61.1	35.4	27.1	90	44.6	23.6	18.1
60	53.3	31.5	24.0	100	40.4	22.1	16.9

Para el nivel de significancia en la prueba del 5 % se ajustó una ecuación polinomial de orden tres (Campos-Aranda, 2003); con los ocho datos de la Tabla 2 se obtuvo la expresión siguiente:

$$PC = \mathrm{C0} - \mathrm{C1} \cdot n + \mathrm{C2} \cdot n^2 - \mathrm{C3} \cdot n^3 \qquad (34)$$

siendo:

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

C0 = 76.20922

C1 = 1.226496

C2 = $9.610048 \cdot 10^{-3}$

C3 = $2.757887 \cdot 10^{-5}$

Test de Wald-Wolfowitz

Esta prueba no paramétrica ha sido utilizada por Bobée y Ashkar (1991), Rao y Hamed (2000), y Meylan *et al*. (2012) para probar *independencia* y *estacionariedad* en registros de gastos máximos anuales (X_i). De acuerdo con el tipo de información hidrométrica procesada, se propuso aplicar la prueba a los registros de gastos pico y volúmenes anuales, que deben ser muestras de valores aleatorios. Wald y Wolfowitz (1943), basándose en el trabajo de Anderson (1942) sobre el coeficiente de correlación serial desarrollaron tal prueba, cuyo estadístico es:

$$R = \sum_{i=1}^{n-1} x_i \cdot x_{i+1} + x_n \cdot x_1 \qquad (35)$$

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Cuando el tamaño (n) de la serie o muestra (x_i) no es pequeño y sus datos son independientes, R procede de una distribución Normal con media y varianza, dadas por las expresiones siguientes:

$$E[R] = \bar{R} = \frac{S_1^2 - S_2}{n-1} \tag{36}$$

$$Var[R] = \frac{S_2^2 - S_4}{n-1} + \frac{S_1^4 - 4 \cdot S_1^2 \cdot S_2 + 4 \cdot S_1 \cdot S_3 + S_2^2 - 2 \cdot S_4}{(n-1)(n-2)} - \bar{R}^2 \tag{37}$$

en las cuales:

$$S_k = \sum_{i=1}^{n} X_i^k \tag{38}$$

Finalmente, se calcula U con la ecuación:

$$U = \frac{R - \bar{R}}{\sqrt{Var[R]}} \tag{39}$$

El valor de U sigue una distribución Normal (0, 1) y se puede usar para probar la independencia de los datos de la serie con un nivel de significancia a, por lo común del 5 %. En una prueba de dos colas, la variable normal estandarizada es $Z_{a/2} \cong 1.96$; entonces, cuando el valor absoluto de U es menor que 1.96, la serie estará integrada por valores independientes (muestra aleatoria).

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Resultados y su discusión

Predicciones univariadas

Con base en los estadísticos del final de la Tabla 1 se aplicaron las ecuaciones (5) y (6) para estimar los parámetros de ubicación y escala de las distribuciones marginales Gumbel (expresiones (3) y (4)), cuyas fórmulas obtenidas son:

$$x_p = 967.390 + 883.361\{-\ln[-\ln(p)]\} \tag{40}$$

$$y_p = 102.736 + 96.974\{-\ln[-\ln(p)]\} \tag{41}$$

De la Ecuación (21) se obtiene:

$$p = 1 - \frac{1}{T_X} \tag{42}$$

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

en la cual T_X es el periodo de retorno en años, igual para las variables X y Y. En la Tabla 3 se muestran las predicciones estimadas con las tres ecuaciones anteriores y los ocho periodos de retorno indicados.

Tabla 3. Predicciones univariadas de la distribución marginal Gumbel en los registros de crecientes anuales de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*), México.

Variable aleatoria	**Periodos de retorno en años**							
	10	**25**	**50**	**100**	**500**	**1 000**	**5 000**	**10 000**
Qp (m^3/s)	2 955	3 793	4 414	5 031	6 456	7 069	8 491	9 103
V (Mm^3)	321	413	481	549	705	773	929	996

Pruebas de las marginales

Para el caso específico de los registros de la Tabla 1, con la Ecuación (31) se obtuvo $|Z| = 2.695$ para *Qp* y $|Z| = 1.436$ para *V*. Entonces, los valores anuales del gasto pico estrictamente no aceptan la distribución Gumbel y los del volumen sí se pueden representar por el modelo Gumbel. Sin embargo, conviene notar que para el *Qp*, el valor absoluto de *Z* no excede de forma notable el valor de 1.645, como lo hace en registros con valores

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

extremos, cuyo valor de k se acerca a -0.55 y aún en registros cortos con n=30, el valor absoluto de Z excede a 4.00.

Con base en las ecuaciones (40) y (41) se calcularon los valores de M_i, para estimar r_x y r_y con las ecuaciones (32) y (33). Se obtuvo 0.9588 para el gasto pico y 0.9912 para los volúmenes. Por otra parte, aplicando la Ecuación (34) con n = 61 se obtuvo un PC = 30.9, por lo cual el r_c = 0.9691. Entonces, el registro de gasto pico no procede estrictamente de una distribución Gumbel ($r_x < r_c$) y el de volúmenes si ($r_y > r_c$).

Verificación de la aleatoriedad

A los registros de gasto pico y volumen anual de la Tabla 1 se les aplicaron las ecuaciones (35) a (39) para probar su independencia y estacionariedad. Se encontró que ambas series son *aleatorias*, con los valores siguientes para el estadístico U: 0.147 y 1.189.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Aceptación gráfica de las marginales

Primero, los gastos pico y volúmenes de la Tabla 1 se ordenaron de menor a mayor. Después, se calcularon sus probabilidades de no excedencia *teóricas* con las ecuaciones (3) y (4), utilizando los parámetros de ubicación y escala de las expresiones (40) y (41). Las probabilidades de no excedencia *empíricas* de ambas series (*Qp* y *V*) se estimaron con la Ecuación (28).

En la Figura 2 y Figura 3 se muestra el contraste gráfico de probabilidades para cada serie ordenada. Se observa un ajuste mucho mejor en el registro de volúmenes. Las diferencias máximas absolutas entre probabilidades empíricas y teóricas de los gastos pico y volúmenes fueron 0.123 y 0.062; la primera ocurrió en el dato ordenado número 49 y la segunda en el 31. Como ambas diferencias son menores que D_n = 0.174 obtenido con la Ecuación (30), se acepta que los registros de *Qp* y *V* de la Tabla 1 siguen la distribución Gumbel.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

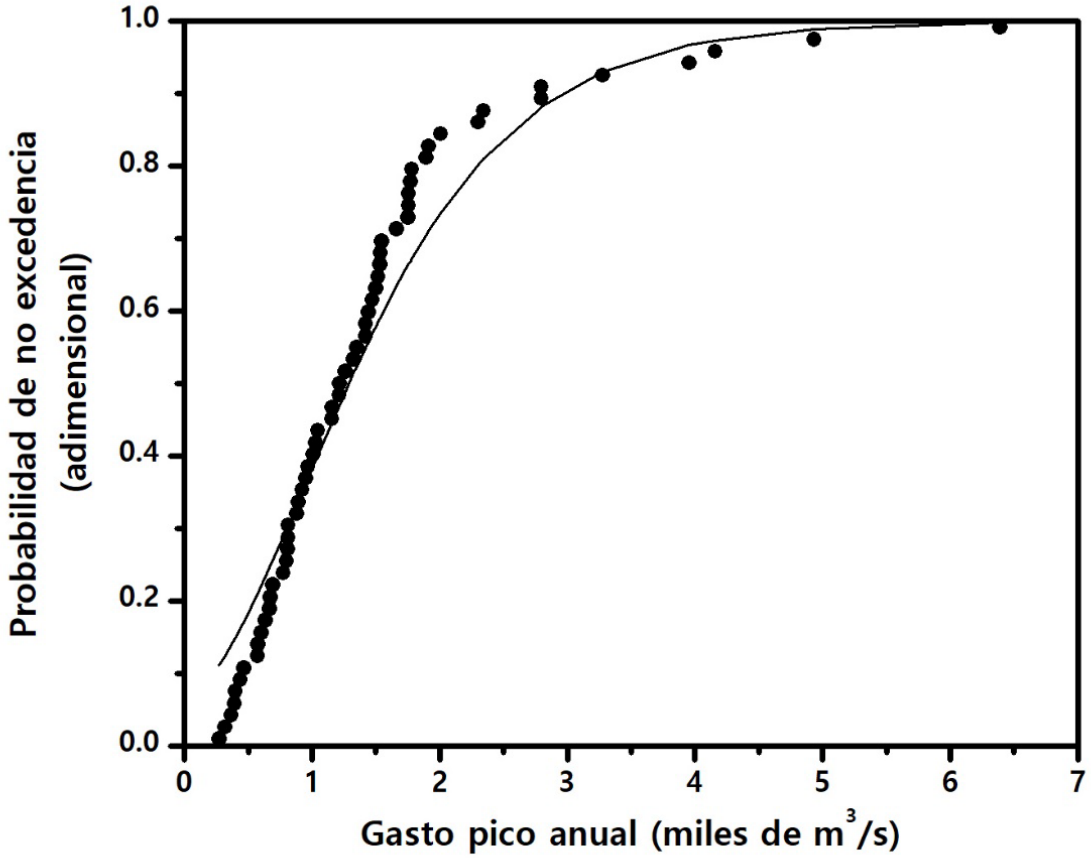

Figura 2. Distribución marginal Gumbel del gasto pico anual de las crecientes entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*), México.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

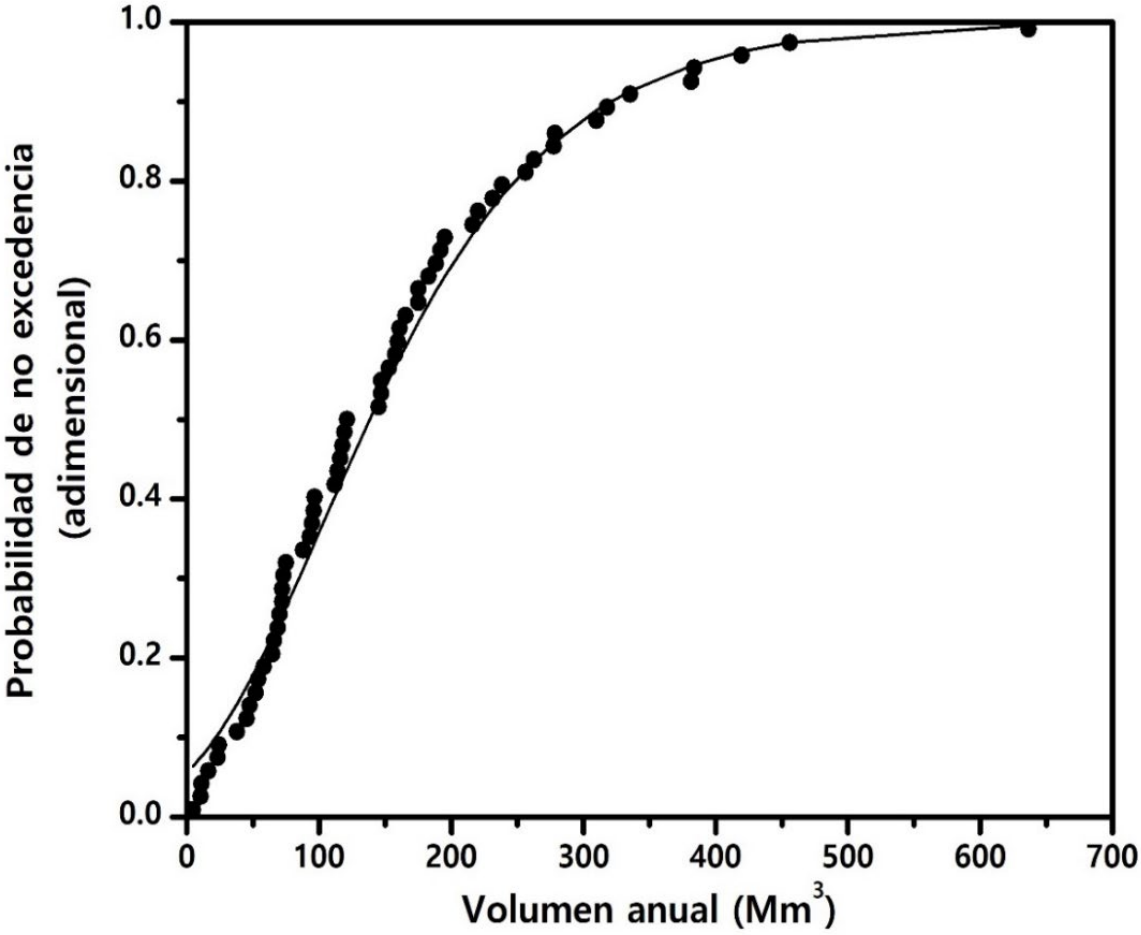

Figura 3. Distribución marginal Gumbel del volumen anual de las crecientes de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*), México.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Contraste de probabilidades conjuntas

La estimación de la probabilidad de no excedencia conjunta *teórica* se llevó a cabo aplicando la Ecuación (2), previo cálculo de la correlación (ρ) con base en la Ecuación (11), entre las variables *Qp* y *V*, cuyo valor de 0.8942 condujo a un parámetro de asociación *m* de 3.0744, según la Ecuación (10).

En cambio, la probabilidad de no excedencia conjunta *empírica* se estimó por medio de la Ecuación (29) y su procedimiento gráfico descrito. En la Tabla 4 se muestran los resultados de ambas probabilidades conjuntas y sus diferencias (empírica menos teórica); también se indican sombreadas las diferencias máximas negativa y positiva. De éstas se deduce que *dif* = 0.1007.

Tabla 4. Probabilidades de no excedencia conjuntas y sus diferencias para las crecientes anuales de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*), México.

Núm.	***F*(*x*,*y*) teórica**	***F*(*x*,*y*) empírica**	**Diferencias**	**Núm.**	***F*(*x*,*y*) teórica**	***F*(*x*,*y*) empírica**	**Diferencias**
1	0.1531	0.1401	-0.0130	32	0.4945	0.5000	0.0055
2	0.1436	0.1237	-0.0199	33	0.4733	0.5000	0.0267
3	0.0981	0.0910	-0.0071	34	0.3735	0.4673	0.0938

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Núm.	*F*(*x*,*y*) teórica	*F*(*x*,*y*) empírica	Diferencias	Núm.	*F*(*x*,*y*) teórica	*F*(*x*,*y*) empírica	Diferencias
4	0.0888	0.0746	-0.0142	35	0.1797	0.1891	0.0094
5	0.0627	0.0419	-0.0208	36	0.2697	0.3364	0.0667
6	0.2789	0.3200	0.0411	37	0.3188	0.4018	0.0830
7	0.0522	0.0092	-0.0430	38	0.7712	0.8109	0.0397
8	0.2396	0.3037	0.0641	39	0.6017	0.6145	0.0128
9	0.9958	0.9908	-0.0050	40	0.4321	0.4673	0.0352
10	0.0938	0.0746	-0.0192	41	0.5535	0.5818	0.0283
11	0.0493	0.0092	-0.0401	42	0.4931	0.5164	0.0233
12	0.1196	0.0910	-0.0286	43	0.8786	0.8927	0.0141
13	0.0544	0.0092	-0.0452	44	0.7883	0.8436	0.0553
14	0.0923	0.0092	-0.0831	45	0.5591	0.5818	0.0227
15	0.3824	0.4673	0.0849	46	0.4004	0.4346	0.0342
16	0.1225	0.0910	-0.0315	47	0.1311	0.1073	-0.0238
17	0.5062	0.5000	-0.0062	48	0.5121	0.5491	0.0370
18	0.3107	0.3855	0.0748	49	0.1416	0.1237	-0.0179
19	0.5713	0.5818	0.0105	50	0.9589	0.9581	-0.0008
20	0.9417	0.9254	-0.0163	51	0.5641	0.5818	0.0177
21	0.5758	0.6145	0.0387	52	0.4303	0.5000	0.0697
22	0.2596	0.3200	0.0604	53	0.2591	0.3200	0.0609
23	0.5723	0.5818	0.0095	54	0.5541	0.5818	0.0277

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Núm.	*F*(*x*,*y*) teórica	*F*(*x*,*y*) empírica	Diferencias	Núm.	*F*(*x*,*y*) teórica	*F*(*x*,*y*) empírica	Diferencias
24	0.4450	0.4509	0.0059	55	0.8884	0.8927	0.0043
25	0.5333	0.5327	-0.0006	56	0.5334	0.5491	0.0157
26	0.5053	0.5000	-0.0053	57	0.5230	0.5164	-0.0066
27	0.6611	0.7618	0.1007	58	0.1990	0.2709	0.0719
28	0.2648	0.2709	0.0061	59	0.7245	0.8109	0.0864
29	0.1990	0.2709	0.0719	60	0.1939	0.2382	0.0443
30	0.2698	0.3364	0.0666	61	0.1749	0.2055	0.0306
31	0.9736	0.9745	0.0009	-	-	-	-

Validación del modelo probabilístico

En la Figura 4 se han dibujado ambas probabilidades de no excedencia conjuntas (teóricas y empíricas de la Tabla 4), observándose un predominio de diferencias positivas, es decir, de puntos por debajo de la recta a 45°. El coeficiente de correlación resultó de 0.990 y el valor del estadístico del test de Kolmogórov-Smirnov es 0.1739 (Ecuación (30)), por lo cual el modelo logístico es aceptado como la distribución conjunta de los datos de la Tabla 1, ya que $dif = 0.1007 < D_n = 0.1739$.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

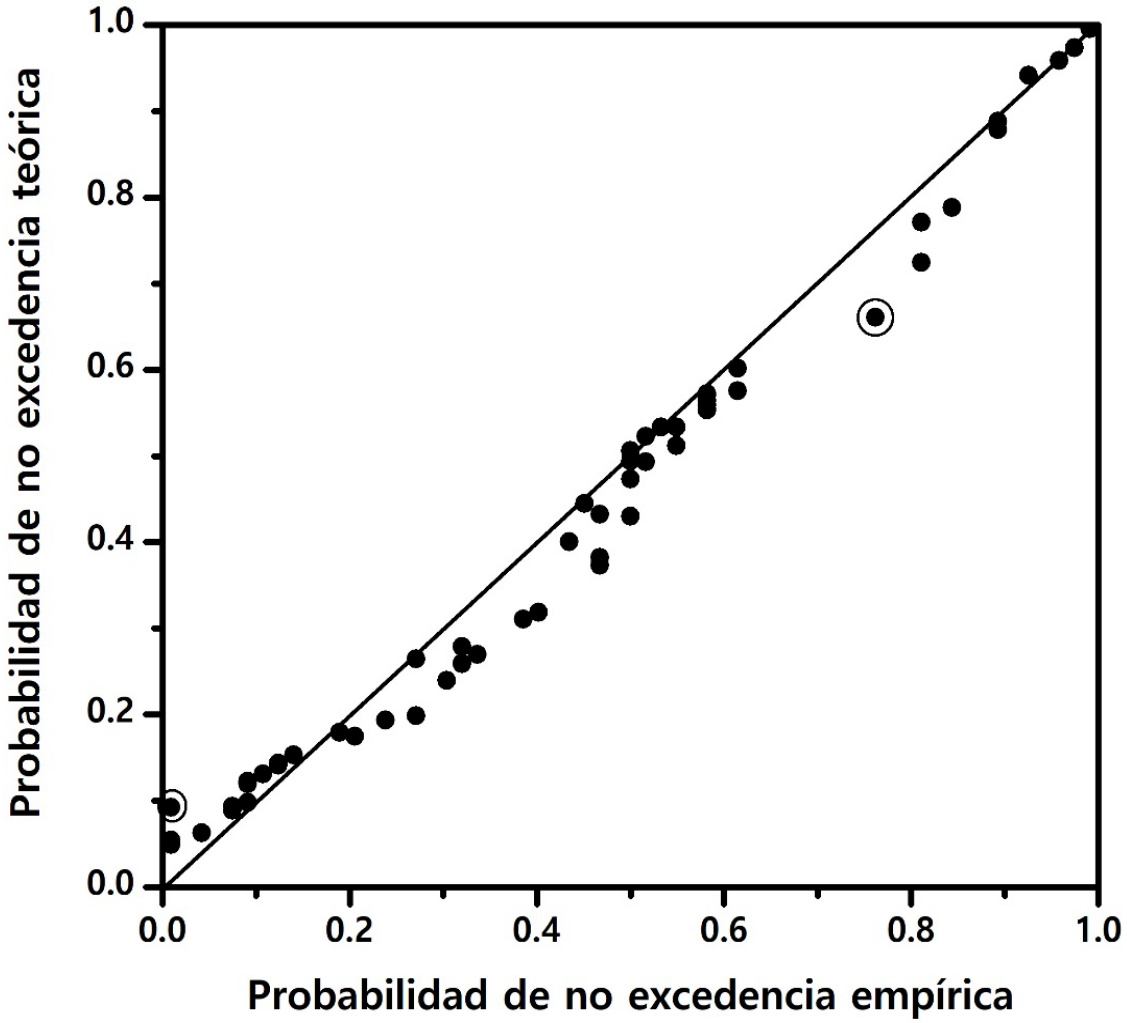

Figura 4. Contraste gráfico de probabilidades conjuntas del gasto pico y volúmenes de las crecientes anuales de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*), México.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Verificación de la restricción de probabilidades

Antes de proceder a estimar los periodos de retorno conjuntos de diseño $T'(Q,V)$, es conveniente verificar la Ecuación (12), que establece las restricciones de la probabilidad. Lo anterior se muestra en la Tabla 5 para un número reducido de parejas de datos históricos.

Tabla 5. Verificación de la restricción de probabilidades del gasto pico y volumen anuales de las crecientes de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*), México.

1	2	3	4	5	6	7
Núm.	***Qp***	***V***	***F*(*x*)**	***F*(*y*)**	***F*(*x*)·*F*(*y*)**	***F*(*x*,*y*)**
1	602.00	64.55	0.220	0.227	0.050	0.153
10	437.00	37.99	0.162	0.142	0.023	0.094
20	3 955.00	381.46	0.967	0.945	0.914	0.942
30	967.00	94.66	0.368	0.337	0.124	0.270
40	1 217.30	159.71	0.471	0.574	0.270	0.432
50	4 158.90	419.64	0.973	0.965	0.937	0.959
61	692.50	68.91	0.255	0.242	0.062	0.175

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Se verificó, en la Tabla 5 completa, que los valores de la columna 7 son mayores que los de la 6, y menores que el más pequeño de las columnas 4 o 5.

Eventos de diseño para obras hidráulicas

Suponiendo que en las cercanías de la entrada a la presa Adolfo Ruiz Cortines y sobre el río Mayo se van a construir diques para protección de planicies de inundación y un puente para cruzarlo, entonces es necesario estimar eventos de diseño con periodos de retorno conjuntos de 500 y 1 000 años. Además, se realizará una revisión de la seguridad hidrológica del tal embalse con periodos de retorno conjuntos de 5 000 y 10 000 años. Por lo anterior, se deben estimar gastos pico y volúmenes anuales con los cuatro periodos de retorno conjuntos citados.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Eventos de diseño obtenidos con regresión

El diagrama de dispersión de las 61 parejas de datos originales (Tabla 1) mostró una nube de puntos con tendencia lineal con un coeficiente de correlación lineal (r_{xy}) de 0.8942. La ecuación de regresión lineal que la representa es la siguiente (Campos-Aranda, 2003):

$$V = 13.6962 + 0.09816 \cdot Qp \tag{43}$$

De la Tabla 3 se obtienen las predicciones siguientes para el gasto pico (*Qp*) y los cuatro periodos de retorno conjuntos de diseño: 6 456, 7 069, 8 491 y 9 103 m^3/s, respectivamente. Con base en la Ecuación (43) se definen los volúmenes anuales siguientes: 647, 708, 847 y 907 Mm^3 para las crecientes de diseño conjuntas.

Al igual que Serinaldi y Grimaldi (2011), se encuentra una gran similitud entre los volúmenes estimados con regresión y sus predicciones de la Tabla 3, que son 705, 773, 929 y 996 Mm^3; es decir, del orden de un 9.5 % mayores.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Eventos de diseño condicionales tipo *T*(*V*|*Q*)

Están definidos por la Ecuación (23). La aplicación de tal fórmula usa la Ecuación (14) y las ecuaciones auxiliares (15) a (20). Para los cuatro periodos de retorno conjuntos de diseño definidos se obtienen de la Tabla 3 los siguientes cuatro gastos pico: 6 456, 7 069, 8 491 y 9 103 m^3/s. Adoptando tales gastos como valores *condicionantes* ($X = x$), se procedió por tanteos del volumen (y) a estimar, con la Ecuación (14), la respectiva probabilidad de no excedencia condicional y con ella en la Ecuación (23), el periodo de retorno condicional, que debe igualar al del gasto pico. Los volúmenes estimados fueron 911, 978, 1 185 y 1 274 Mm^3.

Gráficas del periodo de retorno conjunto *T'*(*Q*,*V*)

Los periodos de retorno conjuntos de tipo AND se estiman con base en la Ecuación (25). Definidos los cuatro periodos de retorno conjuntos de diseño, se seleccionan de forma arbitraria gastos pico y volúmenes, para obtener sus probabilidades de no excedencia marginales y conjunta. Las primeras se estiman con las ecuaciones (3) y (4), y la segunda con la Ecuación (2) del modelo logístico.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

En la Tabla 6 se muestran las parejas de gasto pico y volumen anual utilizadas para definir las cuatro gráficas de la Figura 1 relativas a los periodos de retorno conjuntos de diseño de la aplicación numérica descrita.

Tabla 6. Parejas de gasto pico y volumen anual utilizadas para definir las gráficas del periodo de retorno conjunto (Figura 1) en las crecientes de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*), México.

***T'(Q,V)* 500 años**		***T'(Q,V)* 1000 años**		***T'(Q,V)* 5000 años**		***T'(Q,V)* 10000 años**	
***Vol.* Mm³**	***Qp* m³/s**	***Vol.* Mm³**	***Qp* m³/s**	***Vol.* Mm³**	***Qp* m³/s**	***Vol.* Mm³**	***Qp* m³/s**
400	6 456	400	7 069	400	8 491	400	9 103
500	6 453	550	7 065	500	8 491	800	9 099
525	6 450	600	7 062	650	8 490	825	9 095
550	6 446	625	7 056	700	8 489	850	9 091
575	6 438	650	7 047	750	8 485	875	9 081
600	6 423	675	7 030	800	8 471	900	9 062
625	6 397	700	6 997	825	8 457	910	9 051
650	6 344	725	6 929	850	8 429	920	9 038
675	6 217	740	6 847	875	8 373	930	9 018
680	6 170	750	6 751	880	8 356	940	8 992

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

T'(Q,V) 500 años		*T'(Q,V)* 1000 años		*T'(Q,V)* 5000 años		*T'(Q,V)* 10000 años	
Vol. **Mm³**	***Qp*** **m³/s**	***Vol.*** **Mm³**	***Qp*** **m³/s**	***Vol.*** **Mm³**	***Qp*** **m³/s**	***Vol.*** **Mm³**	***Qp*** **m³/s**
685	6 108	760	6 570	890	8 308	950	8 956
690	6 022	765	6 392	900	8 237	960	8 904
695	5 890	770	5 976	905	8 185	970	8 823
700	5 648	772	5 340	910	8 118	980	8 679
705	4 000	773	4 000	915	8 020	985	8 552
				920	7 864	990	8 332
				923	7 706	993	8 049
				925	7 537	995	7 595
				929	4 000	996	4 000

Selección de los eventos de diseño

En la Tabla 7 se muestran los gastos pico y volúmenes anuales definidos con los periodos de retorno univariados híbridos. Tales parejas de *Qp* y *V* muestran la amplitud de variación que puede tener el volumen anual,

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

pues se definieron adoptando como variable dominante o preponderante al gasto pico.

Tabla 7. Parejas de eventos de diseño obtenidas con los periodos de retorno univariados híbridos en las crecientes de entrada a la presa Adolfo Ruiz Cortines (*Mocúzari*), México.

Evento de diseño	**Periodo de retorno conjunto de diseño, en años**							
	500		**1 000**		**5 000**		**10 000**	
	Qp	***V***	***Qp***	***V***	***Qp***	***V***	***Qp***	***V***
Con regresión	6 456	647	7 069	708	8 491	847	9 103	907
Condicional	6 456	911	7 069	978	8 491	1 185	9 103	1 274

Por otra parte, en la Figura 1 o en la Tabla 6 se pueden seleccionar muchas parejas de *Qp* y *V*, que satisfacen el periodo de retorno conjunto de diseño. En relación con lo anterior, Volpi y Fiori (2012) definen como *subgrupo de parejas críticas* las que están dentro de la curva de cada gráfica de $T'(Q,V)$; es decir, las que no están definidas en las rectas asíntotas.

Desde la **Introdución** se indicó que el AFC bivariado conduce, a través de la Ecuación (25) del $T'(Q,V)$, a una infinidad de combinaciones de gasto pico y volumen que generan el mismo *periodo de retorno*

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

conjunto, y por ello existen muchas crecientes o *hidrogramas* que producirán diferentes efectos en el embalse que se diseña o revisa, adoptando por seguridad el que genera las condiciones más críticas, severas o desfavorables. Lo anterior está incorporando en el diseño hidrológico las características físicas del vertedor y vaso del embalse en proyecto o bajo revisión.

Para formar cada hidrograma de diseño existen métodos teóricos y empíricos (Ramírez-Orozco & Aldama, 2000; Serinaldi & Grimaldi, 2011). Campos-Aranda (2008) ha expuesto un proceso empírico que define hidrogramas tipo Gamma esbeltos y aplanados.

Conclusiones

Se expuso la aplicación del *modelo logístico* (ML) o distribución Gumbel bivariada en el análisis de frecuencias conjunto de los 61 gastos pico y volúmenes anuales de las crecientes que ingresan a la presa Adolfo Ruiz Cortines (*Mocúzari*), en Sonora, México. El ajuste del ML permite el cálculo de las probabilidades univariadas, conjuntas y condicionales. Esta parte del proceso concluye con el contraste de las funciones marginales y la validación del ML, previa estimación gráfica de las probabilidades empíricas univariadas y bivariadas.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Después se abordó el cálculo de los periodos de retorno univariados híbridos y conjuntos. Los primeros utilizan una variable de diseño dominante, por ejemplo, el gasto pico, y definen por regresión el volumen asociado, o lo estiman mediante probabilidad condicional (ecuaciones (14) y (23)). Estas parejas de gasto pico y volumen que se definen son únicas para cada periodo de retorno de diseño. Por otra parte, la búsqueda de las infinitas parejas de gasto pico y volumen (*Qp* y *V*) que generan el mismo *periodo de retorno conjunto* con la Ecuación (25) permiten construir su gráfica, mostrada en la Figura 1.

En el AFC bivariado se definen decenas de parejas de *Qp* y *V*, que generarán *hidrogramas críticos* que producirán diferentes efectos en el embalse que se diseña o revisa, adoptando por seguridad el que genera las condiciones más severas o desfavorables. De esta manera se están incorporando en el diseño hidrológico las características físicas del embalse y su vertedor bajo estudio o revisión.

El ML es una distribución Gumbel bivariada adecuada para procesar de manera simple y conjunta registros de gasto pico y volúmenes de crecientes que no presentan valores extremos dispersos (*outliers*) ni poblaciones mezcladas. En la limitación citada radica su universalidad para procesar crecientes de cuencas medianas y pequeñas de regiones con mecanismos meteorológicos únicos de formación de las crecientes.

Por su naturaleza explícita, el ML es sumamente simple de aplicar y conduce a resultados de importancia en los análisis de frecuencias conjuntos. Debido a su limitación de tener marginales Gumbel, sería conveniente identificar en qué regiones del país los registros de *Qp* y *V*

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

de sus crecientes anuales pueden ser modelados por tal FDP y así tener una idea de su aplicación potencial en México.

Agradecimientos

Se agradecen las observaciones y correcciones sugeridas por los árbitros anónimos B y D, y del editor, las cuales permitieron evitar omisiones de antecedentes y de teoría en relación con los análisis de frecuencias bivariados. Tales correcciones también ayudaron a definir los alcances del estudio y sus posibles líneas de investigación futuras.

Referencias

Aldama, A. A. (2000). Hidrología de avenidas. Conferencia Enzo Levi 1998. *Ingeniería Hidráulica en México*, 15(3), 5-46.

Aldama, A. A., Ramírez, A. I., Aparicio, J., Mejía-Zermeño, R., & Ortega-Gil, G. E. (2006). *Seguridad hidrológica de las presas en México.* Jiutepec, México: Instituto Mexicano de Tecnología del Agua.

Anderson, R L. (1942). Distribution of the serial correlation coefficient. *The Annals of Mathematical Statistics*, 13(1), 1-13. DOI: 10.1214/aoms/1177731638

Bobée, B., & Ashkar, F. (1991). Chapter 1. Data requirements for hydrologic frequency analysis. In: *The Gamma Family and derived distributions applied in Hydrology* (pp. 1-12). Littleton, USA: Water Resources Publications.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Campos-Aranda, D. F. (2003). Capítulo 5. Ajuste de curvas. En: *Introducción a los métodos numéricos: software en Basic y aplicaciones en hidrología superficial* (pp. 93-127). San Luis Potosí, México: Editorial Universitaria Potosina.

Campos-Aranda, D. F. (2008). Procedimiento para revisión (sin hidrometría) de la seguridad hidrológica de presas pequeñas para riego. *Agrociencia*, 42(5), 551-563.

Chow, V. T. (1964). Statistical and probability analysis of hydrologic data. Section 8-I: Frequency analysis (pp. 8.1-8.42). In: Chow, V. T. (ed.). *Handbook of applied hydrology*. New York, USA: McGraw-Hill Book Co.

Domínguez, R., & Arganis, M. L. (2012). Validation of method to estimate design discharge flow for dam spillways with large regulating capacity. *Hydrological Sciences Journal*, 57(3), 460-478. DOI: 10.1080/02626667.2012.665993

Escalante-Sandoval, C. A., & Reyes-Chávez, L. (2002). Capítulo 9. Análisis conjunto de eventos hidrológicos. En: *Técnicas estadísticas en hidrología* (pp. 203-246). México, DF, México: Facultad de Ingeniería, Universidad Nacional Autónoma de México.

Escalante-Sandoval, C. A. (2005). Análisis de eficiencia de la distribución Bi-Gumbel. *Ingeniería. Investigación y Tecnología*, 6(1), 13-17.

Goel, N. K., Seth, S. M., & Chandra, S. (1998). Multivariate modeling of flood flows. *Journal of Hydraulic Engineering*, 124(2), 146-155.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Gräler, B., Van-den-Berg, M. J., Vandenberghe, S., Petroselli, A., Grimaldi, S., De-Baets, B., & Verhoest, N. E. C. (2013). Multivariate return periods in hydrology: A critical and practical review focusing on synthetic design hydrograph estimation. *Hydrology and Earth System Sciences*, 17(4), 1281-1296. DOI: 10.5194/hess-17-1281-2013

Hosking, J. R., & Wallis, J. R. (1997). Appendix: *L*-moments for some specific distributions. In: *Regional frequency analysis. An approach based on L-moments* (pp. 191-209). Cambridge, UK: Cambridge University Press.

Kite, G. W. (1977). Chapter 8. Type I extreme distribution and Chapter 12. Comparison of frequency distributions In: *Frequency and risk analyses in hydrology* (pp. 87-104, 156-168). Fort Collins, USA: Water Resources Publications.

Meylan, P., Favre, A. C., & Musy, A. (2012). Chapter 3. Selecting and checking data series. In: *Predictive hydrology. A frequency analysis approach* (pp. 29-70). Boca Raton, USA: CRC Press.

Ramírez-Orozco, A. I., & Aldama, A. A. (2000). Capítulo 1. Teoría estadística y análisis de frecuencias conjunto. En: *Análisis de frecuencias conjunto para estimación de avenidas de diseño* (pp. 25-58). Avances en Hidráulica No. 7. México, DF, México: Asociación Mexicana de Hidráulica-Instituto Mexicano de Tecnología del Agua.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Rao, A. R., & Hamed, K. H. (2000). Theme 1.8. Tests on hydrologic data (pp. 12-21) and Theme 7.2. The Extreme Value type I EV(1) distribution. In: *Flood frequency analysis* (pp. 229-241). Boca Raton, USA: CRC Press.

Requena, A. I., Mediero, L., & Garrote, L. (2013). A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation. *Hydrology and Earth System Sciences, 17*(8), 3023-3038. DOI: 10.5194/hess-17-3023-2013

Salas, J. D., Obeysekera, J., & Vogel, R. M. (2018). Techniques for assessing water infrastructure for nonstationary extreme events: A review. *Hydrological Sciences Journal*, 63(3), 325-352. DOI: 10.1080/02626667.2018.1426858

Serinaldi, F., & Grimaldi, S. (2011). Synthetic design hydrographs based on distribution functions with finite support. *Journal of Hydrologic Engineering*, 16(5), 434-446. DOI: 10.1061/(ASCE)HE.1943-5584.0000339

Shiau, J. T. (2003). Return period of bivariate distributed extreme hydrological events. *Stochastic Environmental Research and Risk Assessment*, 17(1-2), 42-57. DOI: 10.1007/s00477-003-0125-9

Stedinger, J. R. (2017). Flood frequency analysis. In: Singh, V. P. (ed.). *Handbook of applied hydrology* (2nd ed.) (pp. 76.1-76.8). New York, USA: McGraw-Hill Education.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou, E. (1993). Chapter 18. Frequency Analysis of Extreme Events. In: Maidment, D. R. (ed.). *Handbook of hydrology* (pp. 18.1-18.66). New York, USA: McGraw-Hill, Inc.

Vogel, R. M. (1986). The probability plot correlation coefficient test for Normal, lognormal and Gumbel distributional hypotheses. *Water Resources Research*, 22(4), 587-590.

Vogel, R. M., & Castellarin, A. (2017). Risk, reliability, and return periods and hydrologic design. In: Singh, V. P. (ed.). *Handbook of applied hydrology* (2nd ed.) (pp. 78.1-78.10). New York, USA: McGraw-Hill Education.

Volpi, E., & Fiori, A. (2012). Design event selection in bivariate hydrological frequency analysis. *Hydrological Sciences Journal*, 57(8), 1506-1515. DOI: 10.1080/02626667.2012.726357

Wald, A. & Wolfowitz, J. (1943). An exact test for randomnesss in the non–parametric case based on serial correlation coefficient. *The Annals of Mathematical Statistics*, 14(4), 378–388. DOI: 10.1214/aoms/1177731358

Yue, S. (1999). Applying bivariate Normal distribution to flood frequency analysis. *Water International*, 24(3), 248-254.

Yue, S., Ouarda, T. B. M. J., Bobée, B., Legendre, P., & Bruneau, P. (1999). The Gumbel mixed model for flood frequency analysis. *Journal of Hydrology*, 226(1-2), 88-100.

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Yue, S. (2000a). Joint probability distribution of annual maximum storm peaks and amounts as represented by daily rainfalls. *Hydrological Sciences Journal*, 45(2), 315-326. DOI: 10.1080/02626660009492327

Yue, S. (2000b). The Gumbel mixed model applied to storm frequency analysis. *Water Resources Management*, 14(5), 377-389.

Yue, S. (2000c). The bivariate lognormal distribution to model a multivariate flood episode. *Hydrological Processes*, 14(14), 2575-2588.

Yue, S., & Rasmussen, P. (2002). Bivariate frequency analysis: Discussion of some useful concepts in hydrological application. *Hydrological Processes*, 16(14), 2881-2898. DOI: 10.1002/HYP.1185

Yue, S., & Wang, C. Y. (2004). A comparison of two bivariate extreme value distributions. *Stochastic Environmental Research and Risk Assessment*, 18(2), 61-66. DOI: 10.1007/s00477-003-0124-x

2023, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)